CSC249/449 HW3, Kefu Zhu

3.1 Optimization of DCF:

Prove Equation (3) is the optimal solution for Equation (2)

Equation (2)

$$\epsilon = \frac{1}{MN}(|| \sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^* - \mathrm{Y}||^2 + \lambda \sum_{l=1}^{D} ||W_l||^2)$$

Equation (3)

$$W_l = \frac{\Phi_l(\mathrm{x}) \odot \mathrm{Y}^*}{\sum_{k=1}^{D} \Phi_k(\mathrm{x}) \odot \Phi_k^*(\mathrm{x}) + \lambda}$$

Answer:

From Equation (2), we can get

$$MN \cdot \epsilon = tr\ [(\sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^* - \mathrm{Y})^H \cdot (\sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^* - \mathrm{Y})] + \lambda \sum_{l=1}^{D} tr(W_l^H W_l)$$

Since MN is a constant, so the original loss function ϵ is proportional to $MN \cdot \epsilon$. We can then write

$$\epsilon \propto tr\ [(\sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^* - \mathrm{Y})^H \cdot (\sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^* - \mathrm{Y})] + \lambda \sum_{l=1}^{D} tr(W_l^H W_l)$$

$$\propto tr\ [(\sum_{l=1}^{D} \Phi_l(\mathrm{x})^H \odot W_l^T) \cdot (\sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^*)] - tr[Y(\sum_{l=1}^{D} \Phi_l(\mathrm{x})^H \odot W_l^T)] - tr[Y^H(\sum_{l=1}^{D} \Phi_l(\mathrm{x}) \odot W_l^*)] + tr(Y^H Y) + \lambda \sum_{l=1}^{D} tr(W_l^H W_l)$$

To obtain the optimial value of W_l^*, we calculate the derivative of loss function with respect to W_l^* and set it to zero

$$\frac{\partial \epsilon}{\partial W_l^*} = 0 = [\sum_{l=1}^{D} \Phi_l(\mathrm{x})^H \odot W_l^T \cdot \Phi_l(\mathrm{x})] - Y^* \Phi_l(\mathrm{x}) + \lambda W_l = [\sum_{l=1}^{D} \Phi_l(\mathrm{x})^H \odot \Phi_l(\mathrm{x}) \cdot W_l] + \lambda W_l - \Phi_l(\mathrm{x}) Y^*$$

By rearranging the equation, we can get Equation (3)

$$W_l = \frac{\Phi_l(\mathrm{x}) \odot \mathrm{Y}^*}{\sum_{k=1}^{D} \Phi_k(\mathrm{x}) \odot \Phi_k^*(\mathrm{x}) + \lambda}$$

3.2 Proving Parseval's theorem for 2-d DFT

First, we will try to prove Equation (24)

Equation (24)

$$(D_N \otimes D_M)^H (D_N \otimes D_M) = MNI$$

$\because$

$$D_N \otimes D_M = \begin{bmatrix} D_M^{11} D_N & \dots & D_M^{m1} D_N \\ \vdots & \ddots & \vdots \\ D_M^{1m} D_N & \dots & D_M^{mm} D_N \end{bmatrix}, \quad D_N^H \otimes D_M^H = \begin{bmatrix} (D_M^{11})^H D_N^H & \dots & (D_M^{m1})^H D_N^H \\ \vdots & \ddots & \vdots \\ (D_M^{1m})^H D_N^H & \dots & (D_M^{mm})^H D_N^H \end{bmatrix}$$

$\therefore$

$$(D_N \otimes D_M)^H (D_N \otimes D_M) = (D_N^H \otimes D_M^H)(D_N \otimes D_M)$$

$$= \begin{bmatrix} (D_M^{11})^H D_N^H & \dots & (D_M^{m1})^H D_N^H \\ \vdots & \ddots & \vdots \\ (D_M^{1m})^H D_N^H & \dots & (D_M^{mm})^H D_N^H \end{bmatrix} \begin{bmatrix} D_M^{11} D_N & \dots & D_M^{m1} D_N \\ \vdots & \ddots & \vdots \\ D_M^{1m} D_N & \dots & D_M^{mm} D_N \end{bmatrix}$$

$$= \begin{bmatrix} ((D_M^{11})^H D_M^{11} + \dots) D_N^H D_N & \dots & ((D_M^{m1})^H D_M^{1m} + \dots)^H D_N^H D_N \\ \vdots & \ddots & \vdots \\ ((D_M^{1m})^H D_M^{m1} + \dots) D_N^H D_N & \dots & ((D_M^{mm})^H D_M^{mm} + \dots) D_N^H D_N \end{bmatrix}$$

$$= (D_M^H D_M) \otimes (D_N^H D_N) = (MI) \otimes (NI) = MNI$$

Hence, we can use $vec(X) = (D_N \otimes D_M) vec(X)$ to write $vec(x)$ as

$$vec(x) = \mathcal{F}^{-1}(vec(X)) = (D_N \otimes D_M)^{-1} vec(X) = \frac{(D_N^H \otimes D_M^H) vec(X)}{MN}$$

We can then expand $\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |x[m,n]|^2$ as the following

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |x[m,n]|^2 = vec(x)^T vec(x)$$

$$= \left(\frac{(D_N^H \otimes D_M^H) vec(X)}{MN}\right)^T \left(\frac{(D_N^H \otimes D_M^H) vec(X)}{MN}\right)$$

$$= \frac{1}{M^2 N^2} \cdot (vec(X)^T vec(X)) \cdot [(D_N^H \otimes D_M^H)(D_N \otimes D_M)]$$

$$= \frac{1}{M^2 N^2} \cdot (vec(X)^T vec(X)) \cdot MN$$

$$= \frac{1}{MN} \cdot (vec(X)^T vec(X))$$

$$\because \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |X[m,n]|^2 = \frac{1}{MN} \cdot (vec(X)^T vec(X))$$

$$\therefore \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |x[m,n]|^2 = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |X[m,n]|^2$$